THE PRIVACY ADVISOR

iapp

Data protection issues for employers to consider when using generative AI

By IAPP Member Contributor Zoe Argento, CIPP/US

9 August 2023

The recent explosion of generative artificial intelligence tools coincides with a parallel explosion in privacy legislation, both in the U.S. and around the world. In the U.S., 13 states passed comprehensive data protection laws in less than three years. Globally, most developed countries passed new or stricter privacy laws within the last decade. Many of these laws explicitly regulate the application of AI.

Consequently, feeding personal data into generative AI tools and handling personal data in their outputs entails navigating a thicket of data protection obligations.

These issues are particularly complex with respect to human resources data. Every employer manages troves of personal data about its workforce, much of it highly sensitive, ranging from health information to performance evaluations. For most companies, HR data is the most sensitive data they handle.

There are three key issues in this complex space.

The first concerns disclosure of personal data to AI tools. These disclosures may cause employers to lose control of the data and even result in data breaches. Second, data provided by generative AI services may be based on processing and collecting personal data in violation of data protection requirements, such as notice and the appropriate legal basis. Employers could potentially bear some liability for these violations. Third, when using generative AI services, employers must determine how to comply with requests to exercise data rights in accordance with applicable law.

Risks related to feeding personal data into generative AI

Because generative AI excels in synthesizing and summarizing information, employers may be tempted to use it to produce reports or other products involving HR data. However, submitting personal data to a generative AI tool can put that data at significant risk.

Acme's dilemma

Consider the following hypothetical. Acme Company's CEO requests, on short notice, a PowerPoint presentation on Acme's employee compensation across different divisions and types of employment positions. Eager to please but short on time, the head of HR creates an account on a new generative AI service, uploads key statistics about compensation across its global workforce, including specific compensation for several named individuals, and requests a presentation.

The generative AI service produces several slides with helpful and clear graphics. The head of HR supplements the slides with some she creates herself and provides a presentation on compensation to the CEO the next day. The CEO is pleased, at least that day.

A week later, however, an employee informs the head of HR that he found information about Acme's employees' compensation publicly available on the internet. Word gets around quickly. The employees named in the data are embarrassed and angry that this information leaked. To make matters worse, some employees file a lawsuit for unlawful pay discrimination based on earning disparities revealed by the report.

Of course, Acme's dilemma is a worst-case scenario, but the example illustrates the privacy-related perils of submitting personal data to a generative AI service. As when providing data to any third party, a company must consider issues of security and control.

Disclosure risks

The generative AI service may divulge a user's personal data both inadvertently and by design. As a standard operating procedure, for example, the service may use all information from users to fine-tune how the base model analyzes data and generates responses. The personal data might, as a result, be incorporated into the generative AI tool. The service might even disclose queries to other users so they can see examples of questions submitted to the service.

There may be nothing surreptitious about these practices. Indeed, the terms of use might explain them clearly. As a result, before providing any personal information to a generative AI service, companies should carefully evaluate the terms of use and, if possible, negotiate protections for their data. Under most data protection regimes, the company may be required to execute a data processing agreement with the generative AI service which contains provisions specified by law.

Of course, obtaining contractual assurances is just one step to protecting personal information processed by a generative AI service. A service provider may agree to all the provisions the employer requests, but then suffer a data breach. To reduce this risk, companies should consider conducting due diligence before entrusting their personal information to a generative AI service.

Deidentification

Deidentifying the data before submitting it to a generative AI service can reduce the risk. In contrast to personal data, deidentified data is largely unregulated by privacy laws.

However, most data protection regimes, such as the California Privacy Rights Act and the EU General Data Protection Regulation, set high standards for deidentification. Consequently, simply removing names and identification numbers does not necessarily meet the deidentification standards set by applicable law. The CPRA, for example, requires the business to ensure the recipient of the deidentified data agrees by contract not to reidentify the data.

Risks related to collecting and processing of input data

In addition to the risks of submitting company information to a generative AI tool, there are also risks for employers using data collected by a generative AI service.

As an example, a software company decides to use generative AI to get a better sense of the qualifications of engineers in the potential hiring pool. The company requests a report from a generative AI service on the education and certifications of software engineers in the city where it is located.

In responding to the request, the generative AI service draws from data scraped from the internet about local software engineers. It then produces a report containing a summary, as well as examples of specific individuals.

In this use case, the data protection violations may result from the way the generative AI service obtains the information and generates the report. Critically for employers, an employer may inherit legal risks from the AI service due to the way it handles the input data.

The generative AI service itself would be most directly liable for data protection violations, but the employer also could potentially be liable for using the reports and other outputs from the generative AI service. This is more likely to be the case if the employer signed a service provider agreement in which the generative AI service acts as the agent of the employer. In that case, the generative AI services' actions might be attributable to the employer.

The key ways in which the collection and processing of input data might result in data protection law violations are lack of lawful basis for the processing, failure to provide notice regarding collection and noncompliant transfer of personal data across national borders.

Lawful basis

Most data protection laws around the world permit the collection and processing of personal data on limited grounds only, such as the individual's consent or as required by law. Under these laws, a company cannot scrape data from the internet and use it without a lawful basis.

Depending on the country, establishing the lawful basis may prove to be a substantial hurdle.

For example, in some countries, such as South Korea where lawful processing is heavily based on consent, collecting a large training set might be impractical due to the difficulty of obtaining consent from each individual. Other countries, such as Brazil and members of the EU, permit data processing based on the company's "legitimate interests" to the extent that the individual's rights and freedoms do not outweigh the company's legitimate interests.

However, it is not yet clear whether regulatory authorities will find companies to have legitimate interest in scraping massive amounts of personal data from the internet to train AI systems. EU regulators, for example, have publicly expressed concerns. In fact, Italy's data protection authority, the Garante, briefly banned generative AI service ChatGPT in part because of the lack of a lawful basis for its data collection.

In the U.S., data protection laws generally have not adopted the concept of a lawful basis for processing personal data, but this could change as more states pass data protection laws.

Notice

Virtually all data protection laws require the organization that collects and processes personal data to provide a notice regarding how it processes the personal data. These notices must be detailed, including descriptions of the purposes of using the personal data and the parties to which it is disclosed.

In some circumstances, the notice must provide details about how the algorithm works. For example, under the GDPR, employers must provide notice to employees about decisions made solely based upon AI that produce "legal effects concerning him or her or similarly significantly affects him or her." This likely includes decisions about hiring and termination of employment.

In that case, the notice must provide "meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject." Given that most AI tools maintain their algorithms as trade secrets, an employer might not be able to provide this information when using a generative AI tool.

As a practical matter, meeting a requirement to provide notices to the individuals at issue may prove an insuperable hurdle if the AI service simply copies vast amounts of personal data from the internet. Applicable laws typically require notice at the point of collection. In the case of mass data scraping, there may be no feasible way to contact or even identify the individuals at issue.

Data protection laws in the U.S. are different than most such laws in that they generally carve out publicly available information from their definitions of personal data. As a result, scraping data on the internet may not qualify as the collection of personal data because the data is publicly available.

Nevertheless, companies should beware of the nuance in definitions of publicly available under U.S. state data protection laws.

Data is only publicly available if lawfully made available through government records or widely distributed media, or the individual lawfully made that information available. If the information became available without the individual's knowledge or consent, as in the hypothetical of the payroll records discussed above, then the information may still be "personal data" and fully protected under the applicable state data protection law.

Companies should consider steps to ascertain that the collected data is publicly available. For example, to increase the likelihood that the individual consented to public posting of their information, they might ensure the generative AI service only collects data from websites with user-created bios and secure user accounts.

Finally, a company may be required to provide its own notice to individuals about processing their personal data. If an employer requests a report about an individual from a generative AI service, such as a report on an applicant based on what the service finds about them online, then many data protection laws would require the employer to provide its own notice about the collection and use of the report to the individual.

Cross-border data transfers

Most data protection laws prohibit the transfer of personal data to another country except in limited circumstances. If the generative AI service collects personal data from one country and transfers it to an employer in another country, the transfer may violate the first country's data protection laws. Employers should evaluate the data flow and adopt lawful data transfer mechanisms to address this issue and avoid potential violations.

This can be complicated because the lawful data transfer mechanisms vary both by sender and recipient country. Many countries have identified a limited list of countries that provide adequate data protection. Data transfers to these adequate countries may not entail any compliance hurdles.

However, transfers to other countries might require consent of the individual, which may be impractical. Still other countries, such as the EU member states, allow data transfers without consent, but only if a derogation applies or a lawful data transfer mechanism, such as standard contractual clauses, has been implemented.

Implications of privacy risks in collecting and processing AI input data

Given the many risks of using personal data collected by AI tools, employers should vet the tools they use and negotiate service agreements to reduce their risk. In these service agreements, employers should consider requiring assurances that the generative AI service complied with applicable data protection laws when collecting and processing personal data.

As a backstop, employers might require indemnification provisions and stipulate that the generative AI service maintains substantial liability insurance for data-related claims.

Risks related to the collection and processing of input data

The issues of lawful basis, notice, and cross-border data transfers largely relate to how the personal data is collected and transferred. A company's use and retention of the personal data raises additional privacy issues.

Right to delete

Depending on the applicable data protection laws, individuals may have rights to access, delete, correct or stop the processing of their personal data. The right to delete poses a particular challenge to generative AI.

What if an employee demands the deletion of personal data that a company already submitted to an AI platform to develop a custom tool or to fine-tune the AI? Depending on the nature of the generative AI system, the system may not be capable of truly "forgetting" data points. The AI "learns" by recognizing patterns in the training data and using those patterns and data points to generate new content. As a result, personal data may be embedded in the AI's patterns.

Accuracy

Generative AI may produce inaccurate content for several reasons. The training data set, as well as source data, may be wrong. Also, the tool itself may simply make up information, the so-called "hallucinations" produced by some AI tools.

Most data protection laws provide a right to correct personal data, at least to the extent that it is inaccurate. Correcting data in the training data set raises similar issues to deleting data.

In addition, outside of the U.S., data protection laws nearly universally require data controllers to ensure the accuracy of personal data. As a result, an employer could be liable under these laws for not vetting generative AI tools for accuracy and relying on inaccurate reports generated by these tools.

In the U.S., the new data protection laws generally do not impose this accuracy obligation, though they do require organizations to correct inaccurate information upon request. The Fair Credit Reporting Act, however, requires accuracy and creates a separate set of risks.

The FCRA regulates "consumer reporting agencies" and people who use the reports generated by consumer reporting agencies. Crucially, a generative AI service potentially could meet the definition of "consumer reporting agency" if the service regularly produces reports about individuals' "character, general reputation, personal characteristics, or mode of living" and these reports are used for employment purposes.

The FCRA is a hyper-technical, high-risk law. It requires employers to provide disclosures and obtain authorization before using reports from consumer reporting agencies for employment purposes.

Among other points, the FCRA obligates employers to notify an individual before taking adverse action based on a report, provide a copy of the report and disclose contact information for the consumer reporting agency.

Due to its private right of action and statutory damages, class action lawsuits alleging violations of the FCRA often result in six-figure settlements.

Conclusion

Before providing personal data to a generative AI service or using personal data from the service, employers should think through the data protection implications.

Key considerations are the source and nature of the data, applicable data protection laws and purposes for using the information. Depending on these factors, employers may need to provide notices, obtain consent, aquire detailed contractual assurances from the service provider and implement processes for handling data rights requests.

iapp CPE credits approved Approved

CDPO, CDPO/BR, CDPO/FR, CIPM, CIPP/A, CIPP/C, CIPP/E, CIPP/G, CIPP/US, CIPT, LGPD

Credits: 1

SUBMIT FOR CPES